BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 3: e34826
doi: 10.3897/biss.3.34826

OPEN ACCESS

Conference Abstract

Uncertain Geo-Uncertainty: Ecological research and public biodiversity data repositories

Arnald Marcer[‡], Elspeth M. Haston[§], Quentin Groom[|], F. Xavier Picó[¶], Agustí Escobar[‡], Francesc Uribe[#]

‡ CREAF, E08193 Bellaterra (Cerdanyola del Vallès), Catalonia, Spain
§ Royal Botanic Garden Edinburgh, Edinburgh, United Kingdom
| Meise Botanic Garden, Meise, Belgium
¶ Departamento de Ecología Integrativa, Estación Biológica de Doñana (EBD), Consejo Superior de Investigaciones Científicas (CSIC), Sevilla 41092, Spain
# Museu de Ciències Naturals de Barcelona, Barcelona, Spain

Corresponding author: Arnald Marcer (arnald.marcer@uab.cat)

Received: 25 Mar 2019 | Published: 13 Jun 2019

Citation: Marcer A, Haston E, Groom Q, Picó F, Escobar A, Uribe F (2019) Uncertain Geo-Uncertainty: Ecological research and public biodiversity data repositories. Biodiversity Information Science and Standards 3: e34826. https://doi.org/10.3897/biss.3.34826

Abstract

Natural history collections represent a vast and superb wealth of information gathered and curated across centuries by institutions such as natural history museums and botanical gardens around the world. The relatively recent advent and maturation of accessible computer technology has allowed the initiation of major digitization projects aimed at making the contents of these collections publicly available for education and research purposes. The final destinations of these newly digitized data are public biodiversity data repositories, of which, GBIF is the main one. These respositories are gateways where researchers can access and retrieve the data for use in a wide range of analyses.

This unprecedented volume of information on biodiversity represents an extraordinary asset for research in ecology and evolution. A particularly important part of the digitized data for any given specimen is its collection location, as it indirectly gives information on the species' habitat and thus, its ecological requirements. Many specimens in natural history collections come from a time where the collecting event, which includes the location information, was hand-written on physical tags attached to the specimen. This location

© Marcer A et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

information was given as a description of a place, e.g. a site name, and could be a rather precise or vague description. In order to convert this description of locality into a digitized research-grade georeferenced record, the research community has come up with a set of guidelines and recommendations; the most prominent one the point-radius method devised by Wieczorek et al. in 2004. However, and despite the public availability of this know-how, the end result is that the data available at the end of the pipeline, e.g. GBIF, often lacks georeferencing information with enough quality to be used for research purposes. Occurrence records from natural history collection datasets held at GBIF, often lack spatial coordinates and, if present, in most cases their precision and uncertainty fields are blank. The final consequence of this lack of complete georeferencing information is that the affected records are rendered useless for many kinds of research. For example, the flourishing field of species distribution modelling absolutely depends on accurate spatial information in order to be able to retrieve information on the environmental conditions in which the species live. The availability of global environmental and remote sensing datasets together with the sophisticated geospatial tools at the disposal of the researcher become powerless if no quality geoinformation is available.

In this study, we perform a preliminary analysis on the status and availability of geoferencing information in datasets originated from specimens in natural history collections held at GBIF, discuss how the quality of this spatial info may affect ecological research, and conclude with some recommendations on how to better describe the georeferencing process within public digital biodiversity repositories.

Keywords

public data repositories, georeferencing, data flows, uncertainty, ecological research

Presenting author

Arnald Marcer